In the claims:

1. (Currently amended) A method of detecting speech in an incoming signal comprising the steps of:

receiving said incoming signal, ~~performing a preprocessing step of~~ extracting an ~~noise~~ estimate of the noise background of the incoming signal and suppressing the noise background of the incoming signal to provide a noise suppressed signal in which the estimated background noise has been removed ~~to augment signal-to-noise ratio of a speech signal~~, filtering the noise suppressed signal in which the background noise has been removed with a spectral inverse filter, said spectral inverse filter is determined by spectrum maxima and the inverse filtering operation comprising the steps of :

in the logarithmic (dB) domain, removing the mean spectral magnitude from the original speech spectrum,

in the mean removed short term frequency spectrum *S(i)*, (*i*=1…128), determining all the frequency position (*Pj*), whose magnitudes are maxima over a window centered around *Pj* and stretching N positions to the left and right of *Pj*,

in the list of peaks, adding the first (i=1) and last (i=128) frequency positions, their associated magnitudes set equal to the mean of the first and last M x N magnitudes, respectively,

removing the mean of the peak magnitudes from each peak magnitude,

if the largest resulting peak magnitude exceeds MAX_dB_DN, normalizing all peaks so that the largest peaks magnitude becomes MAX_dB_DN, and

the resulting inverse filtering *H(i)*, (i=1…128) is defined as the maximum of the normalized peaks and 0 dB, and

removing the inverse filter from the original spectrum in the logarithmic domain $U(i) = S(i) - H(i)$ and measuring the periodicity of the ~~incoming~~ signal from the inverse filter using an autocorrelation function to determine whether a signal frame correspond to a speech frame or not.[[.]]

2. (Original) The method of claim 1 wherein said periodicity measurement is defined as :

$$\rho = \max_{T_l}^{T_h} Rx(\tau)$$

where T_l and T_h are pre-specified so that the period will range in the range of speech and the signal is speech if ρ is above a given threshold.

3. (Original) The method of Claim 2 wherein said period is between about 75 Hz and 400 Hz.

4. (Currently amended) The method of claim [[1]] 2 where said threshold value is set to maximize speech detection accuracy.

5. (Original) The method of claim 1 wherein said extracting step includes the steps of:

converting the spectrum of the incoming signal into logarithmic domain,

removing high frequency components in logarithmic domain by recurrent filtering along the time axis,

establishing an estimate of noise background, converting the estimate into linear domain, and

suppressing the noise background from the signal, in linear domain.

6. (Canceled)

7. (Currently amended) The method of Claim ~~6~~ 10 wherein said inverse filtering is based on a normalized approximation of the envelope of the short term speech spectrum derived from a local maxima of the short term speech spectrum.

8. (Currently amended) The method of claim [[8]] 7 wherein said inverse filtering is performed in a log frequency domain and is implemented by subtracting from the original spectrum the estimated inverse filtering spectrum.

9. (Canceled)

10. (New) A noise-resistant utterance detector comprising the steps of:

accepting a speech utterance input signal,

removing background noise from the utterance signal according to a spectral subtraction method to get a noise subtracted signal,

filtering the noise subtracted signal with a spectral inverse filter to get an inverse filtered signal,

locating close low-frequency formants in the noise subtracted signal if they exist and inserting spectral valleys between said formants before inverse filtering,

calculating the autocorrelation from the inverse filtered signal to get an autocorrelation result, and

detecting that a frame of the signal being processed is or is not speech based on a threshold applied to the autocorrelation result.

11. (New) The method of claim 10 wherein said spectral inverse filter is determined by spectrum maxima and the inverse filtering operation by the steps of :

in the logarithmic (dB) domain, removing the mean spectral magnitude from the original speech spectrum,

in the mean removed short term frequency spectrum *S(i)*, (*i*=1...128), determining all the frequency position (*Pj*), whose magnitudes are maxima over a window centered around *Pj* and stretching N positions to the left and right of *Pj*,

in the list of peaks, adding the first (i=1) and last (i=128) frequency positions, their associated magnitudes set equal to the mean of the first and last M x N magnitudes, respectively,

removing the mean of the peak magnitudes from each peak magnitude,

if the largest resulting peak magnitude exceeds MAX_dB_DN, normalizing all peaks so that the largest peaks magnitude becomes MAX_dB_DN,

the resulting inverse filtering *H(i)*, (i=1...128) is defined as the maximum of the normalized peaks and 0 dB, and

removing the inverse filter from the original spectrum in the logarithmic domain $U(i)= S(i)- H(i)$.

In the Specification:

Please amend the specification as follows:

[0030] First, the short term speech spectrum of the speech frame is normalized, with a mean equal to zero dB. Then, a battery of tests is performed to detect the presence of two close low-frequency formants. If we determine the following parameters,

σ_1 : The relative magnitude of the first estimated formant,

σ_2 : The relative magnitude of the second estimated formant,

λ_1 : Index in the frequency axis (1...128) of the first estimated formant,

λ_2 : Index in the frequency axis (1...128) of the first estimated formant,

a flag signaling the presence of two close low-frequency formants is raised if the following conditions are met:

1. $\sigma_1 \geq \tau_1$, $\sigma_2 \geq \tau_2$ and $(\sigma_1 - \sigma_2) \leq \tau$,
2. $\lambda_1 \geq \lambda_{min}$ and $\lambda_1 \leq \lambda_{max}$,
3. $(\lambda_2 - \lambda_1) \geq \delta_{min}$ and $(\lambda_2 - \lambda_1) \leq \delta_{max}$.